

Evolutionary branching in a stochastic population model with discrete mutational steps

S. Sagitov^a, B. Mehlig^{b,*}, P. Jagers^a, V. Vatutin^c

^aMathematical Sciences, Chalmers and Gothenburg University, SE-41296 Gothenburg, Sweden.

^bDepartment of Physics, University of Gothenburg, SE-41296 Gothenburg, Sweden.

^cDepartment of Discrete Mathematics, Steklov Mathematical Institute, Moscow, Russia

Abstract

Evolutionary branching is analysed in a stochastic, individual-based population model under mutation and selection. In such models, the common assumption is that individual reproduction and life career are characterised by values of a trait, and also by population sizes, and that mutations lead to small changes ϵ in trait value. Then, traditionally, the evolutionary dynamics is studied in the limit $\epsilon \rightarrow 0$. In the present approach, small but non-negligible mutational steps are considered. By means of theoretical analysis in the limit of infinitely large populations, as well as computer simulations, we demonstrate how discrete mutational steps affect the patterns of evolutionary branching. We also argue that the average time to the first branching depends in a sensitive way on both mutational step size and population size.

Keywords: Evolutionary branching, genetic drift, selection, adaptation

1. Introduction

The development of populations subject to frequency- or density-dependent selection is an important and central but not easily analysed topic in theoretical biology. Many different models and approximation schemes have been developed in order to understand how trait values change and how the intriguing possibility of ‘evolutionary branching’ may arise, where a population of individuals with a well-defined trait value or genotype (monomorphic, in other words) experiences a split leading to two or more coexisting trait values or genotypes in the population. This process is important because it gives rise to increased biodiversity.

A variety of population-genetic and game-theoretic methods, as well as the body of ideas and approximation schemes known as *adaptive dynamics*, cf. Metz et al. (1996); Geritz et al. (1998); Dieckman (2004); Waxman and Gavrillets (2005); Metz et al. (2012), have been employed in order to study evolutionary branching. Several simplifications are commonly resorted to. First, complications due to mating and recombination in diploid populations are disregarded by ascribing to each individual a single trait value x that is transferred by clonal reproduction unless mutation occurs (see however the recent interesting approach to adaptive dynamics with Mendelian inheritance by Collet et al., 2011). Selection is taken to act through a trait- and density-dependent fitness function, identified with mean reproduction. Second, it is assumed that fitness is a smooth function of the trait value x and the population density. The third postulate is that the mutation rate μ is so small that only few trait values are represented in the population at any one time. Mutations thus occur so rarely that we can talk

of a separation of time scales between long-term evolutionary dynamics (the sequence of mutations) and short-term population dynamics, defined by birth and death events, an evolutionary versus an ecological time scale. Fourth, it is assumed that mutations lead only to small changes in the trait value (that mutational step sizes ϵ are small). Fifth, the population size N is taken to be large, and effects of random genetic drift are neglected (apart from incorporating the fixation probability of an advantageous mutation).

In the adaptive-dynamics literature a further step is taken by letting $N \rightarrow \infty$, $\mu \rightarrow 0$, and $\epsilon \rightarrow 0$. Then evolution of a monomorphic population to the first branching of traits is analytically described by the so-called ‘canonical equation’ (Dieckmann and Law, 1996), or more generally by deterministic evolution on the fitness landscape (Waxman and Gavrillets, 2005a; Lande, 1976; Iwasa et al., 1991). How these deterministic adaptive dynamics may arise from a strict stochastic population model has been described in detail in the elegant treatment of Champagnat (2006) and later papers.

As a monomorphic (single-trait) population approaches a local fitness maximum where evolutionary branching may occur in the sense that populations with two different trait values may coexist from there onwards, the canonical equation fails to describe the dynamics. It becomes necessary to ask how stochasticity in a finite population affects the possibility of evolutionary branching. Further questions are: what is the shape of the bifurcation diagram of trait values in the wake of a branching? How do the values of μ and ϵ influence the evolutionary dynamics in a finite population? How long does it take (in the mean) for evolutionary branching to occur?

We address these matters analytically and by computer simulations of a stochastic, individual-based model for the evolution of trait values in a finite population subject to density-dependent

*Corresponding author

Email address: Bernhard.Mehlig@physics.gu.se (B. Mehlig)

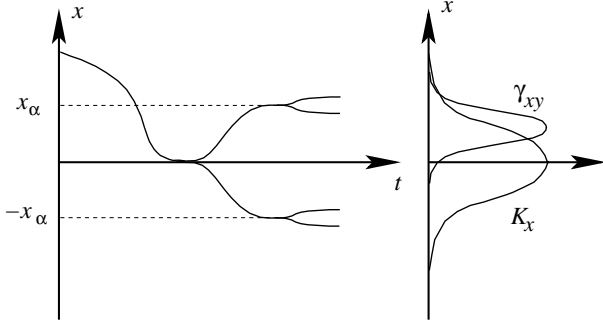


Figure 1: A schematic view of the deterministic model. The left panel depicts the first three evolutionary branching events. The right panel illustrates the main ingredients of the fitness function given by Eq. (2). The function γ_{xy} is plotted for a particular value of $y = x_0$ which in the case shown is positive.

fitness function symmetric with respect to its maximal value at the trait $x = 0$. In the limit of $N \rightarrow \infty$, $\mu \rightarrow 0$, and ϵ small, we show that the evolution of a monomorphic population towards the fitness optimum at $x = 0$ is governed by a canonical equation. For our model, standard stability analysis in the limit of $\epsilon \rightarrow 0$ would predict evolutionary branching of a monomorphic population at $x = 0$. We demonstrate that at small but fixed mutational step sizes ϵ , evolutionary branching typically occurs at non-zero values of x . In the limit of infinite population size and vanishing mutation rate we compute the critical value of x where the first branching is most likely to occur, as a function of ϵ . In addition we show that the evolutionary dynamics in this limit gives rise to a non-symmetric bifurcation diagram after the first branching (for a symmetric model such as ours, the standard theory predicts a symmetric bifurcation diagram in the limit $\epsilon \rightarrow 0$). Moreover, we state geometrical conditions determining the trait values where the second evolutionary branching event is most likely to occur.

Finding the average time until evolutionary branching occurs in our model requires a more refined analysis. It is necessary to specify how $N \rightarrow \infty$ and $\epsilon \rightarrow 0$. This determines the stochastic population dynamics in the critical stage preceding the first branching. That stage must last long enough to render branching possible. We discuss results of computer simulations relating to this question briefly, but a precise mathematical description of the problem is left for future work.

The remainder of this paper is organised as follows. In Section 2 we introduce the stochastic, individual-based model used in simulations. Section 3 summarises standard results obtained in the limit $N \rightarrow \infty$, $\mu \rightarrow 0$, and $\epsilon \rightarrow 0$. These are contrasted with numerical experiments at relevant values of N , μ , and ϵ in Section 4. Discrepancies between the $\epsilon \rightarrow 0$ predictions and the results of numerical experiments are discussed. Our results (and the calculations supporting them) addressing the issues raised in Section 4 are summarised in Section 5. Section 6 contains our conclusions. Three appendices summarise details of our calculations.

2. Model

The stochastic, individual-based model used in the numerical experiments discussed below describes a finite population of individuals competing for resources. Each individual is assigned a trait value x . Individuals with the same trait value (belonging to the same ecological ‘niche’) compete for a limited amount of resources. The scarcity of resources is parametrised by a trait-dependent carrying capacity K_x . There is an ‘optimal’ trait value corresponding to the value of x where K_x assumes its global maximum. Individuals belonging to different niches (corresponding to different trait values, x and y , say) may also interact. The strength of interaction depends upon the difference $|x - y|$ in trait values.

Mutations give rise to changes in the trait value, $x \rightarrow x \pm \epsilon$. Here $\epsilon > 0$ is a fixed mutational step size and the set of possible traits occurring in the population is $x = 0, \pm\epsilon, \pm 2\epsilon, \dots$. Individual fitness is measured by the mean offspring number, which is determined by the trait values and the intensity of competition. The latter, in its turn, is controlled by the scarcity of resources and the number of individuals. In other words, the model describes the evolution of a trait value, or of trait values, in a population subject to density-dependent selection. The population would be expected to evolve towards the optimal trait value. Intense competition between individuals with trait values close to the optimal one may however give rise to sub-populations with well-defined but distinct trait values. This phenomenon is referred to as *evolutionary branching*, and is the subject of this paper.

The number of individuals with a given trait value x is denoted by Z_x . It is assumed that each individual has either none or two offspring (Klebaner et al., 2011). The offspring of individuals in one generation constitute the next generation. We denote the fitness function by $M_x(Z_y, y \in D)$ which is the mean offspring number for an individual with trait value x , and where D is the set of trait values represented in the population. This fitness is a function of the co-existing sub-population sizes $(Z_y, y \in D)$. We take it to be of the form:

$$M_x(Z_y, y \in D) = \frac{2}{1 + (NK_x)^{-1} \sum_{y \in D} \gamma_{xy} Z_y}. \quad (1)$$

Here the carrying capacity of the subpopulation with trait value x is NK_x , where N is a population-size parameter to be thought of as large. Competition between subpopulations corresponding to different trait values x and y is regulated by the function γ_{xy} .

The offspring usually inherit the parental trait x , but with a small probability μ they mutate (independently) either to $x - \epsilon$ or to $x + \epsilon$, with equal probabilities. Since time is measured in generations, the model is thus discrete both in time and trait space. It can be viewed as the simplest possible structure (‘the bare bones’, Klebaner et al., 2011) behind adaptive dynamics, and as such is an interesting mathematical object in its own right. Since all individuals have the same life span, one, it displays an unyielding age dependence which can be seen as a diametrical counterpart and natural complement to the equally artificial non-aging individuals of birth-and-death processes, underlying

the sequence of papers by Champagnat and Méléard (2011) and others, and also classical deterministic approaches. For more realistically age-structured models cf. Méléard and Tran (2009) as well as Jagers and Klebaner (2011).

The form (1) is certainly ad hoc, and should be viewed as an explorative illustration. But it is not without historical precedence (Christiansen and Loeschke, 1980). As there, we take K_x and γ_{xy} to be Gaussian functions. Indeed, this is a choice often encountered in the adaptive-dynamics literature,

$$K_x = e^{-x^2} \quad \text{and} \quad \gamma_{xy} = e^{-(x-y)^2(1+\alpha)}. \quad (2)$$

The parameter α regulates the competition between subpopulations with different traits. A larger value of α implies weaker competition. We take $\alpha > 0$. This ensures that competition between subpopulations is weak enough to allow for evolutionary branchings. It is convenient to replace population sizes Z_x by densities $f_x = Z_x/(NK_x)$, so that the fitness function (1) takes the form

$$M_x(f_y, y \in D) = \frac{2}{1 + \sum_{y \in D} R_{xy} f_y}. \quad (3)$$

Here

$$R_{xy} = \gamma_{xy} K_y / K_x = e^{-(x-y)(\alpha x - (2+\alpha)y)} \quad (4)$$

is the fitness kernel which combines the effects of the interaction γ_{xy} and the carrying capacity K_x .

Eqs. (3) and (4) allow for explicit formulae for the equilibrium densities ($f_y^D, y \in D$) at which subpopulations may coexist in the absence of mutations. Note however that even this coexistence is temporary, albeit potentially of long duration (Jagers and Klebaner, 2011). We shall still refer to them as equilibrium densities. They are defined by the condition $M_x^D = 1$, where

$$M_z^D \equiv \frac{2}{1 + \sum_{y \in D} R_{zy} f_y^D}. \quad (5)$$

The equation $M_x^D = 1$ describes the critical regime of reproduction. It corresponds to the condition

$$\sum_{y \in D} R_{xy} f_y^D = 1, \quad x \in D. \quad (6)$$

In the monomorphic case $D = \{x\}$ we find that $M_x(f_x) = 2/(1 + f_x)$ and in the absence of mutations, the density f_x stabilises at the value $f_x^{(x)} = 1$. Substituting this result into Eq. (5) we obtain

$$M_z^{(x)} = \frac{2}{1 + e^{-(z-x)(\alpha z - (2+\alpha)x)}}. \quad (7)$$

Importantly, the function M_z^D given by Eq. (5) plays a crucial role in the invasion analysis. Suppose an individual of type $z \notin D$ appears in a stable population with the set of traits D as a mutant of $x \in D$. Its fitness has the form

$$M_z(f_y^D, y \in D; f_z^*) = \frac{2}{1 + \sum_{y \in D} R_{zy} f_y^D + f_z^*} \approx M_z^D$$

because $R_{zz} = 1$ and the small initial frequency of the mutant f_z^* can be neglected. If $M_z^D > 1$, there is a chance for the new

trait to establish itself in the set of coexisting traits. If $M_z^D \leq 1$, the mutation can not establish itself and is wiped out by genetic drift.

In summary, we have chosen an extremely simple model which however retains so much of the structure of evolutionary dynamics that the questions raised in the introduction can still be formulated. It is described by only four parameters, the population-size parameter N , the mutation rate μ , the mutational step size ϵ , and the interaction parameter α .

3. Standard predictions in the limit $\epsilon \rightarrow 0$

In spite of its simplicity, the model introduced in the previous Section is not easily analysed. As pointed out, the standard adaptive-dynamics approach (Geritz et al., 1998) studies continuous-time population models in the limit $N \rightarrow \infty$ and $\mu \rightarrow 0$, and then $\epsilon \rightarrow 0$. In this Section we state the corresponding results for our discrete-time model.

If N can be thought of as infinite, law-of-large-number effects replace random processes by their expectations, and if both $\mu \rightarrow 0$ and $N \rightarrow \infty$, then only a small number of trait values are represented in the population at any specific time: the basic cases studied below are those of monomorphic populations, $D = \{x\}$, and dimorphic ones, $D = \{x_1, x_2\}$.

In this limit, a time-scale separation appears between the evolutionary dynamics (referring to long-term changes of the trait value) and the ecological time scale of population dynamics. Letting $\epsilon \rightarrow 0$ makes it possible to describe the dynamics of the trait value x by local stability analysis, central to adaptive dynamics.

Fig. 1 illustrates the standard predictions obtained in these limits. First, starting from an initially positive trait value, $x(t)$ is expected to evolve deterministically towards the local optimum (at $x = 0$ in our model). The path $x(t)$ is described by a differential equation referred to as a ‘canonical equation’ (Dieckmann and Law, 1996; Champagnat et al., 2001) and related to similar equations studied in population genetics (Lande, 1976; Iwasa et al., 1991). Second, the first evolutionary branching occurs at $x = 0$. Third, after the first branching, the pair of traits evolves deterministically in a symmetric fashion, and the second and third branchings occur simultaneously.

3.1. Deterministic evolution towards the first branching

Assume that the population starts from a positive initial trait value $x_0 = j_0 \epsilon$ which is large compared to the small value of the mutational step size ϵ : $j_0 \gg 1$. Classical adaptive dynamics then predicts that the initial trait value is consecutively replaced by smaller values of x , driving $x(t)$ towards $x = 0$. The resulting function of time $x(t)$ is described by a differential equation for $x(t)$, known as the canonical equation (of trait evolution), the standard reference being Dieckmann and Law (1996). In the notation of Champagnat et al. (2001, see Eq. (7a)) the classical canonical equation of trait evolution has the form

$$\frac{dx}{dt} = \mu(x) \frac{\sigma_0^2}{2} n(x) \partial_1 f(x, x). \quad (8)$$

In this expression, $\mu(x)$ is the mutation rate for the trait value x , σ_0^2 is the mutational variance, $n(x)$ is the equilibrium population size, and $f(z, x)$ is the fitness function. Recall that within a model formulated in terms of stochastic birth and death processes, the fitness function $f(z, x)$ is the difference between the birth and death rates for rare mutants with trait z in an x -population at its quasi-stable size.

In our case, the mutation rate $\mu(x) = \mu$ is independent of the trait value x , σ_0^2 is simply ϵ^2 , and $n(x) = Ne^{-x^2}$. The fitness function $f(z, x)$ corresponding to our discrete-time model should be calculated as

$$f(z, x) = \log M_z^{[x]}, \quad (9)$$

where $M_z^{[x]}$ is given by Eq. (7). This formula relies on a well-known property of a linear birth-death process stemming from a single individual: if f is the difference between the birth and death rates per individual, the population size at time t has mean e^{tf} . Eq. (9) stipulates that the expected population size $e^{f(z, x)}$ after one unit of continuous time is equal to the expected size $M_z^{[x]}$ of the first generation in our discrete-time model.

From Eqs. (9) and (7) we conclude that the fitness gradient is given by

$$\partial_1 f(x, x) = -x,$$

since $M_x^{[x]} = 1$. In our case we would thus expect the canonical equation to take the form

$$\frac{dx}{dt} = -\mu \frac{\epsilon^2}{2} Ne^{-x^2} x. \quad (10)$$

However, the correct canonical equation for our model Eq. (19) derived in Section 5.1 reveals that the trait substitution process goes twice as fast in the discrete-time model compared to the continuous-time model. This is reminiscent of the well-known phenomenon of genetic drift running twice as fast in the Moran model than in the Wright-Fisher model of the same size, see for example Wakeley (2008), though the latter phenomenon has a different explanation.

3.2. Location of the first evolutionary branching

The canonical equation suggests that the substitution process of the trait value slows down as the point $x = 0$ is approached. In the limit of $N \rightarrow \infty$, $\mu \rightarrow 0$, and $\epsilon \rightarrow 0$, standard adaptive-dynamics analysis predicts the first evolutionary branching to occur at $x = 0$, giving rise to two branches, $x_1 < 0$ and $x_2 > 0$.

3.3. Second and third evolutionary branchings

Moreover, in the very same limit, the deterministic evolution after the first branching is predicted to be symmetric. In other words, the two trait values x_1 and x_2 after the first branching are expected to evolve as $x_1(t) = -x_2(t)$. The second and third evolutionary branchings are expected to occur simultaneously when $x_2(t)$ reaches a critical value, which in our case can be computed explicitly as

$$x_\alpha = \frac{1}{2} \sqrt{\frac{\log(2\alpha + 1)}{1 + \alpha}}. \quad (11)$$

As a function of α the value of x_α reaches its maximum 0.3731... at $\alpha = 1.2955...$

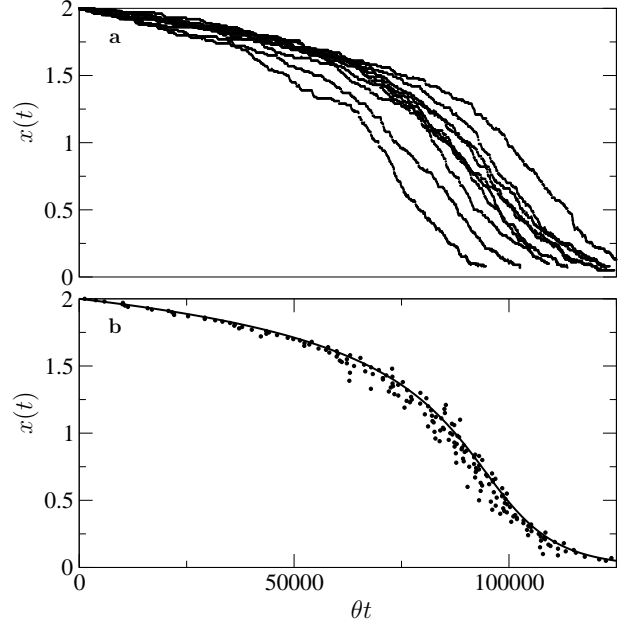


Figure 2: **a** Shows ten realisations of the evolution of $x(t)$ from an initially large value $x_0 = 2$ for a mutational step size of $\epsilon = 10^{-2}$ (additional parameters: $N = 10^6$, $\mu = 10^{-8}$, $\alpha = 9$, $\theta = \mu N = 10^{-2}$). **b** Shows the trait value x as a function of the average time of staying at x , conditional on that no branching has occurred (dots). Also shown is an approximate solution of (10) (solid line).

4. Computer simulations

Figs. 2, 3 and 4 summarise results of direct numerical simulations of the model described in Section 2 for large values of N , small values of μ , and for a small value of ϵ (equal to 10^{-2}). In the algorithm, binomial distributions of numbers of mutants have been approximated by the appropriate Poisson distributions.

We now compare simulation results with the predictions from Section 3. Fig. 2a shows ten realisations of the evolution of $x(t)$ for the mutational step size $\epsilon = 10^{-2}$ from an initially large value $x_0 = 2$. (The additional parameters are given in the figure caption). We see that $x(t)$ decreases towards the first branching (which in all cases occurs for positive values of x and not at $x = 0$). Fig. 2b shows a plot of the trait value x versus the average time of staying at x , conditional on no branching having occurred. Also shown is an approximate solution of the correct canonical equation derived in Section 5.1 as Eq. (19). This is expressed in terms of the expected time for $x(t)$ to reach level x :

$$t(x) = \sum_{y=x-\epsilon}^{x_0} \lambda_y^{-1}, \quad (12)$$

where $\lambda_x = \mu N \epsilon x e^{-x^2}$ (see below) is the rate of the asymptotically exponential holding time at level x . For large values of x we observe good agreement between (12) and the average of ten independent realisations of the substitution process. But as $x = 0$ is approached, the numerical average falls below the predicted average. The neighbourhood of the first branching event is not described by the canonical equation.

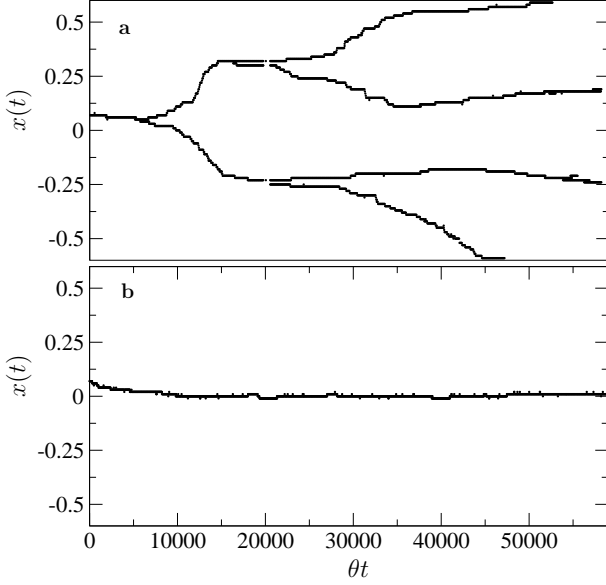


Figure 3: Results of computer simulations of trait evolution in the stochastic individual-based model described in Section 2. Parameters: $\alpha = 9$, $\epsilon = 0.01$, and $\mathbf{a} N = 5 \times 10^6$, $\mu = 2 \times 10^{-10}$.

Fig. 3 shows results of two computer simulations starting in the vicinity of $x = 0$. In Fig. 3a the first branching does not occur at $x = 0$ but at $x = 0.05$. Further, evolution after the first branching is not symmetrical. This can also be seen from Fig. 4 which displays the data of Fig. 3a in the x_1 - x_2 plane (the standard adaptive-dynamics approach predicts, as pointed out above, that the pair $(x_1(t), x_2(t))$ moves on the diagonal shown in Fig. 4 as a dashed line). Note also that the second and third branchings do not occur simultaneously, and not symmetrically in the simulations.

Fig. 3b shows results of a direct numerical simulation with the same overall substitution rate $\theta = \mu N$ as Fig. 3a, but for a smaller value of N and a larger value of μ . In this example, branching did not occur during the simulation time $\theta t < 10^5$. By contrast, the trait value is seen to diffuse around $x = 0$.

In summary we find that the standard adaptive-dynamics approach describes the initial time evolution of the trait value $x(t)$ well, provided $x(t)$ is large enough. In the vicinity of and after the first branching, however, the observed patterns of evolutionary dynamics (Figs. 3 and 4) differ from the standard predictions. In the following we show that the asymmetric patterns observed in Figs. 3a and 4 are due to the positive mutational step size ϵ . In the limit $N \rightarrow \infty$ and $\mu \rightarrow 0$ we were able to characterise the branching patterns at small but fixed values of ϵ . To determine how long it takes on average until evolutionary branching occurs in a finite population requires a more refined analysis. We return to this in Section 5.3.

5. Results

5.1. Deterministic evolution towards the first branching

In the monomorphic case ($D = \{x\}$) Eq. (7) implies

$$M_z^{[x]} - 1 \approx x(x - z) + (\alpha/2 - x^2)(z - x)^2, \quad (13)$$

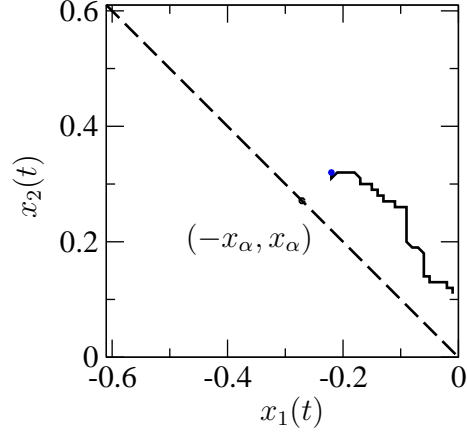


Figure 4: Shows the trajectory of Fig. 3a after the first branching in the x_1 - x_2 -plane. Only points corresponding to $x_1 < 0$ are shown. Parameters the same as Fig. 3a.

for $z \approx x$. For $x > 0$ it follows that

$$M_{x-\epsilon}^{[x]} > 1, M_{x+\epsilon}^{[x]} < 1, M_x^{[x-\epsilon]} < 1, \quad \text{and} \quad M_x^{[x+\epsilon]} > 1. \quad (14)$$

These inequalities show that the expected trend of adaptive evolution in the monomorphic phase is simple: the initial positive trait value is consecutively replaced by smaller values driving $x(t)$ towards zero. The dynamics of this trait substitution process is described by the canonical equation whose correct form for our model is derived next.

Observe that the expected number of mutations $x \rightarrow x - \epsilon$ in the x -population, assumed to be at its quasi-stable size, $Z_x = NK_x = Ne^{-x^2}$, is $(\mu/2)Ne^{-x^2}$ per generation, where μ is divided by 2 since we discard the $x \rightarrow x + \epsilon$ -mutations. Thus the expected waiting time between two consecutive replacements is

$$\lambda_x \approx \frac{\mu}{2} Ne^{-x^2} P_{x-\epsilon}(x). \quad (15)$$

Here $P_z(x)$ stands for the probability that a population from one z -individual survives and takes over from the monomorphic x -population. To find the probability $P_z(x)$ of the mutant $z = x - \epsilon$ invading and replacing a resident population with trait x we can approximate the population dynamics of the invader by a binary branching process starting from a single individual. The survival probability of this branching process is

$$P_z(x) = 2 \frac{M_z^{[x]} - 1}{M_z^{[x]}}, \quad (16)$$

see for example Eq. (5.66) in Haccou et al. (2005). Eq. (13) implies

$$M_{x-\epsilon}^{[x]} \approx 1 + \epsilon x, \quad (17)$$

and we obtain $P_{x-\epsilon}(x) \approx 2\epsilon x$, at least for x far from zero. (The corresponding classical result for the Wright-Fisher model is the following. Suppose that an advantageous allele is introduced in the population with its mean offspring number being larger compared to the wild type by a factor $(1 + s)$. Then the fixation probability of this allele is approximately given by $P \sim 2s$ for small values of s .)

Substituting $P_{x-\epsilon}(x) \approx 2\epsilon x$ into Eq. (15) results in $\lambda_x \approx \mu \epsilon N e^{-x^2} x$. Now, since the evolution of the trait value $x(t)$ as a function of time is approximately given by

$$-\frac{dt}{dx} \approx \frac{t(x-\epsilon) - t(x)}{\epsilon} \approx \frac{1}{\epsilon \lambda_x}, \quad (18)$$

we conclude that the canonical equation takes the form

$$\frac{dx}{dt} = -\mu \epsilon^2 N e^{-x^2} x \quad (19)$$

which is similar (but not identical) to Eq. (10). Indeed, Eqs. (19) and (10) differ by a factor of $\frac{1}{2}$.

This factor can be explained as follows. The standard derivation of Eq. (8) (Dieckmann and Law, 1996; Champagnat et al., 2001) refers to birth-and-death processes in continuous time (or corresponding deterministic formulations), whereas our model considers binary splitting in discrete, non-overlapping generations. Our form of the equation appears explicitly for Poisson reproduction in discrete time in Metz (2011). The fact that different effective population sizes can appear in the canonical equation was first noted by Durinx, Metz, and Meszéna (2008).

Note that the speed of the trait-substitution process given by Eq. (8) would not change if we were to take different values of the birth and death rates, as long as their difference $f(z, x)$ remains the same. We argue that the proper choice of the birth rate in the monomorphic stable population is unity, matching the birth rate in our discrete-time model: one birth per generation per individual, $M_x^{(x)} = 1$. Then, in accordance with Eq. (9), the corresponding survival probability of a single mutant is given by $f(x - \epsilon, x) = \log M_{x-\epsilon}^{(x)} \approx \epsilon x$. Comparing this with its discrete-time counterpart $P_{x-\epsilon}(x) \approx 2\epsilon x$, we conclude that the survival probability in the discrete model is two times larger.

An important assumption behind Eq. (19) is that successful mutations should not come too soon after each other in order to make sure that selective sweeps do not overlap ('clonal interference'), see Su-Chan Park et al. (2010). As a very rough approximation, we estimate the average fixation time from the equation

$$(M_{x-\epsilon}^{(x)})^{T_{\text{fix}}} \approx N. \quad (20)$$

This equation simply suggests to neglect the competition among individuals and equate the target population size N with the mutant reproduction factor for T_{fix} consecutive generations. Combining Eqs. (17) and (20) we find $T_{\text{fix}} \approx \log N / (\epsilon x)$ in the absence of competing mutations. The average time between successful mutations is λ_x^{-1} . So the condition ensuring non-overlapping sweeps is that the product $\lambda_x T_{\text{fix}} \approx \mu N \log N e^{-x^2}$ must be very small so that

$$\mu N \log N \ll 1. \quad (21)$$

In other words, condition Eq. (21) guarantees that after a mutation the new set of types settles in an equilibrium before the next successful mutation event.

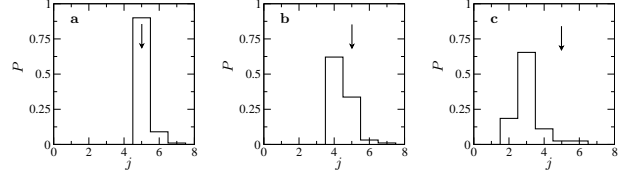


Figure 5: Shows histograms of the locations j of the first branching event. All histograms are normalised to unity. Initial condition $j = 7$. Parameters: $\alpha = 9$, $\epsilon = 10^{-2}$. The first evolutionary branching is expected to occur in the interval $-5 \leq j \leq 5$. The upper limit of this interval is indicated by black arrows. Panel a: $N = 5 \times 10^6$, $\mu = 2 \times 10^{-10}$, all 100 runs resulted in an evolutionary branching. Panel b: $N = 10^6$, $\mu = 10^{-9}$, out of total 100 total only 85 runs resulted in an evolutionary branching while 15 runs had no branching during the simulation time. Panel c: $N = 10^5$, $\mu = 10^{-8}$, out of total 100 total only 81 runs succeeded and 19 had no branching during the simulation time.

5.2. Location of the first evolutionary branching

The trait substitution process considerably slows down as x approaches zero, so that it must be considered as a sequence of discrete jumps of size ϵ towards zero, rather than a continuous process. At levels close to zero a more refined analysis of fitness asymptotics is required.

In the following we derive an approximation for the location of the first evolutionary branching. We show that it happens close to zero, but typically not at $x = 0$. More precisely, we demonstrate that evolutionary branching becomes possible in an interval of width of order ϵ around $x = 0$. It is thus necessary to distinguish between trait values separated by a few mutation steps. For simplicity we assume in the following that initially $x > 0$, so that the trait substitution process approaches $x = 0$ from above (as in the simulation results shown in Fig. 3). The problem of determining the location of the first branching event is symmetric and corresponding expressions for $x < 0$ are easily found.

The aim is to find trait values $x > 0$ where mutations to and from $x - \epsilon$ are mutually invasive. The corresponding condition is: $M_{x-\epsilon}^{(x)} > 1$ and $M_x^{(x-\epsilon)} > 1$. Using Eq. (13) and dropping the terms much smaller than ϵ^2 we arrive at two conditions:

$$\begin{aligned} M_{x-\epsilon}^{(x)} - 1 &\approx x\epsilon + \alpha\epsilon^2/2 > 0, \quad \text{and} \\ M_x^{(x-\epsilon)} - 1 &\approx -(x-\epsilon)\epsilon + \alpha\epsilon^2/2 > 0. \end{aligned} \quad (22)$$

We conclude that the first evolutionary branching is possible for $0 < x < (1 + \alpha/2)\epsilon$, so that populations with trait values x and $x - \epsilon$ are mutually invasive provided:

$$\epsilon \leq x \leq j^* \epsilon \quad \text{with} \quad j^* = \lceil \alpha/2 \rceil. \quad (23)$$

Here $\lceil \alpha/2 \rceil$ denotes the smallest integer greater than $\alpha/2$. (To avoid complications arising in the situation when $j^* = 1 + \alpha/2$ and $M_x^{(x-\epsilon)} - 1 = o(\epsilon^2)$ for $x = j^* \epsilon$, we will assume that α is not an even integer.) The fact that the critical trait value $j^* \epsilon$ is larger for larger values of α is very intuitive: as competition becomes weaker, initially impossible invasion of $(j^* - 1)\epsilon$ into $j^* \epsilon$ becomes favorable once the benefit of reduced competition outweighs the cost of having a "worse" trait value.

Fig. 5 shows where the first evolutionary branchings occurred in an ensemble of 300 simulations of the stochastic,

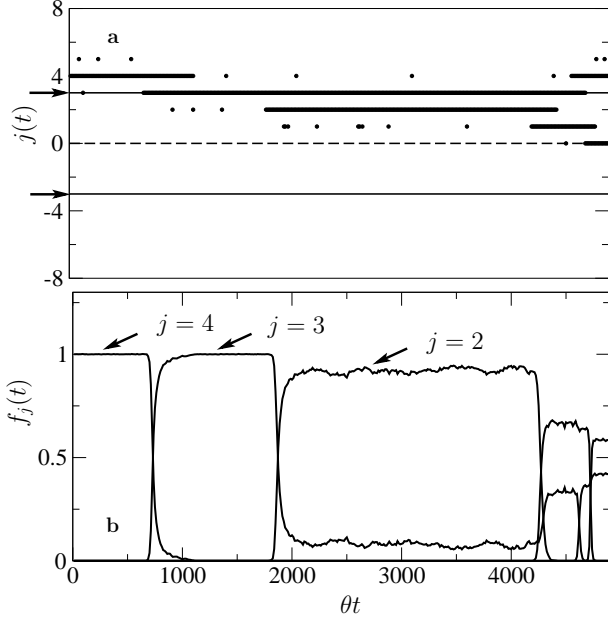


Figure 6: Shows coexistence of trait values prior to the first branching. **a** Evolution of the integer trait value $j(t)$ as a function of time. Parameters: $\alpha = 5$, $\epsilon = 10^{-2}$, $N = 5 \times 10^6$, $\mu = 10^{-9}$. The branching region is given by $-3 \leq j \leq 3$. **b** Dependence of $f_j(t)$ upon time. Clearly seen is the coexistence of the trait values $j = 2, 3$ at $f_2 = 11/12$ and $f_3 = 1/12$. In the case shown, coexistence lasts sufficiently long for evolutionary branching to occur.

individual-based model. The parameter α was chosen to be $\alpha = 9$ which implies that $j^* = 5$. The parameters were chosen so that the overall substitution rate was the same for all runs, $\theta = \mu N = 10^{-3}$. All simulations were started at $x_0 = 7\epsilon$, and were run for the same total time t , given by $\theta t = 10^5$. In a small number of runs evolutionary branching did not occur during this time. Fig. 5 shows where the first branching occurred for the remaining runs. First, we see that apart from a small number of cases, branching occurs in the region predicted. Second, the larger the population size N is, the more likely it is that evolutionary branching occurs at the boundary of the branching region ($j^* = 5$ in this case). To explain this behaviour requires a more refined analysis. We return to this question below in Section 5.3. Third, in the limit $\epsilon \rightarrow 0$ the condition (23) is consistent with the standard prediction in this limit (that the first branching occurs at $x = 0$). Fourth, some branchings occurred one mutation step earlier than expected. This can be explained by the fact that occasionally it can take a long time for a slightly advantageous mutant to stabilise, long enough for the next mutation to initiate the first evolutionary branching.

5.3. Critical state of coexistence prior to first branching

In the preceding section we have derived conditions determining where the first evolutionary branching may occur, Eq. (23). However, results of our computer simulations of the stochastic model described in Section 2 also demonstrate that while evolutionary branching may occur (and often does occur) when these conditions for evolutionary branching are met, branching need not necessarily take place immediately.

This raises the question how long it takes on average in a finite population for evolutionary branching to occur (a question also raised by Boettiger (2010)). What is the probability that evolutionary branching happens when the substitution process reaches the boundary of the region where the first evolutionary branching becomes possible (Fig. 5)? Fig. 6 shows a realisation of a computer simulation of this situation for $\alpha = 5$ (corresponding to $j^* = 3$). In Fig. 6a we see how the trait value reaches the critical value $j^* = 3$ from above, and how subpopulations with trait values $j = 3$ and $j = 2$ coexist until evolutionary branching occurs. The corresponding population sizes f_j are shown in Fig. 6b. For evolutionary branching to occur, the critical state of coexistence must last sufficiently long for a favourable mutation to occur. It is possible, however, that the critical state of coexistence may spontaneously disappear by a fluctuation. We see in Fig. 6b that the population-size fluctuations in the coexistence state are substantially larger than the population-size fluctuations in the monomorphic states.

A preliminary analysis shows that while the fluctuations of the monomorphic population sizes are of order $N^{-1/2}$, the fluctuations of the subpopulation sizes in the critical state of coexistence are much larger: of order $(N\epsilon^2)^{-1/2}$ in the limit of large values of N and small values of ϵ . The smaller value ϵ takes, the shorter is the average life time of the critical state of coexistence, lowering the probability that evolutionary branching occurs in this state. Without going into details, we just briefly sketch our argument (which remains to be made rigorous). Let (f_k, g_k) be the densities $(f_{j\epsilon}^{(j\epsilon, j'\epsilon)}, f_{j'\epsilon}^{(j\epsilon, j'\epsilon)})$ of two populations coexisting at time k and having the neighbouring trait values $j\epsilon$ and $j'\epsilon$ with $j' = j - 1$. The evolution of the vector (f_k, g_k) can be approximated by a stochastic dynamical system

$$\begin{cases} f_{k+1} = \frac{2f_k}{1+f_k+ag_k} + \frac{1}{\sqrt{N}}u_k, & a = e^{-\epsilon^2(\alpha+2-2j)}, \\ g_{k+1} = \frac{2g_k}{1+bf_k+g_k} + \frac{1}{\sqrt{N}}w_k, & b = e^{-\epsilon^2(\alpha+2j)}, \end{cases} \quad (24)$$

where u_k and w_k are independent normal random variables with zero means and standard deviations $\frac{2\sqrt{f_k(f_k+ag_k)}}{1+f_k+ag_k}$ and $\frac{2\sqrt{g_k(bf_k+g_k)}}{1+bf_k+g_k}$ respectively. Provided the deterministic part has a nontrivial stable point we consider a linearised version of this system around this stable point. Our analysis indicates that while the sum $f_k + g_k$ behaves as a stationary autoregression process with fluctuations of the order $N^{-1/2}$, the difference $f_k - g_k$ behaves as a stationary autoregression process with fluctuations of order $(N\epsilon^2)^{-1/2}$. Given that the fluctuations in the sizes of coexisting populations featuring the trait values $(j^*\epsilon, (j^* - 1)\epsilon)$ are of order $(N\epsilon^2)^{-1/2}$ a rough estimate of the probability of the sudden loss of coexistence is obtained using the approximate tail probability for the normal distribution $c_1 e^{-c_2 N\epsilon^2}$, where c_1 and c_2 are positive constants. Thus, the life time T of the critical state of coexistence is expected to be inversely proportional to the probability of sudden loss of coexistence $c_1 e^{-c_2 N\epsilon^2}$ so that

$$\log T \sim S_\alpha N\epsilon^2 \quad (25)$$

for some constant S_α depending on α through (23). Fig. 7 shows simulation results for the life time T as a function of N and ϵ . The results of the simulations are consistent with the

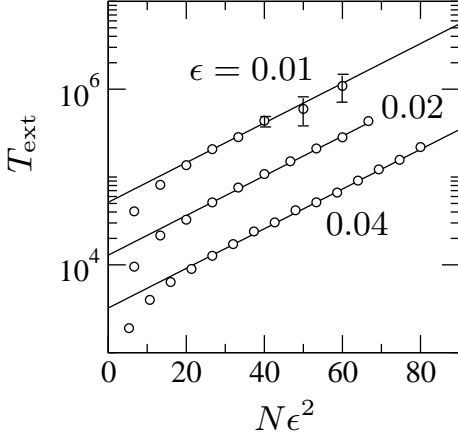


Figure 7: Shows time T_{ext} to extinction of the coexistence state at the boundary of the branching region for the first branching, in the absence of mutations (symbols). Parameters: $\mu = 0$, $\alpha = 3$, $\epsilon = 0.01, 0.02$, and 0.04 . Unless otherwise indicated, two-sigma error bars are smaller than the symbol size. The error bars were determined from 1000 independent simulations, with three exceptions: for $\epsilon = 0.01$ and $N\epsilon^2 = 40, 100$ independent simulations were used, for $N\epsilon^2 = 50$ and 60 , 20 independent simulations. Also shown are fits to the law (25), lines. We find $S_3 \approx 0.052$.

expectation Eq. (25). We find $S_3 \approx 0.052$. The precise mathematical derivation of Eq. (25) and the calculation of the constant S_α are interesting questions for further work.

5.4. Evolution after the first branching

Condition Eq. (23) demonstrates that if the first evolutionary branching occurs, then it happens within a region of size ϵ around $x = 0$. We now turn to the evolution of the pair of trait values (x_1, x_2) after the first evolutionary branching. Standard theory suggests that the dimorphic population evolves symmetrically (along the dashed line in Fig. 4). Our simulations of the stochastic, individual-based population model with discrete mutational steps show that the pair of coexisting trait values followed through the chain of consecutive replacements does not develop in a fully symmetrical fashion. The random path on the x_1 - x_2 -plane (see Fig. 4) follows the trajectory of a random walk with steps from (x_1, x_2) to either $(x_1, x_2 + \epsilon)$ or $(x_1 - \epsilon, x_2)$ depending on which of the two branches the next replacement has been successful. The repulsion between the two branches is due to the pressure of competition. It is advantageous to stay further apart reducing interspecies competition.

Observe that the consecutive steps in this random walk are weakly negatively dependent: the further the walk deviates from the diagonal $x_1 = -x_2$, the larger is the probability for the next step to decrease this deviation. To see this, consider a pair of coexisting populations having trait values x_1 and x_2 with $x_1 < 0 < x_2$ and for definiteness $x_2 > |x_1|$. For a given branch, the corresponding replacement rate is a product of two factors (see the discussion leading to Eq. (15) in the monomorphic case): the stable population size and the probability of fixation for a single mutant. We show that given the lower branch is closer to zero, $x_2 > |x_1|$, both factors of the replacement rate for the lower branch are larger making the move of the lower branch (always off zero, see Appendix B) more probable. In-

deed, as the negative trait value is closer to zero, its stable population size must be larger than the size of the population with the positive trait value. On the other hand, as computed in Appendix B, the probability of fixation for a single mutant in the x_1 -population is also larger, Eq. (49).

The last observation implies that with high probability the locations of the pair of branches satisfy

$$x_1 = -x_2 + O(\sqrt{\epsilon}) \quad (26)$$

prior to the second evolutionary branching. This means that the the expected location of the second evolutionary branching, if any, should not deviate from the diagonal more than a constant times $\sqrt{\epsilon}$. Indeed, the predicted deviation from the diagonal can only be made larger if we neglect the negative dependence and consider a symmetric random walk with independent steps. Referring to the classical central limit theorem for the simple symmetric random walk we observe that the quantity $x_1 + x_2$ (which is always zero in the symmetric case) is of order $\epsilon\sqrt{n}$, where n is the number of jumps in the random walk (representing the number of replacements since the first evolutionary branching). Since the second evolutionary branching takes place at values of x_1 and x_2 of order unity, one must wait for a large number of jumps, namely of order $1/\epsilon$ for the second branching to occur. We conclude that the path taken by the two branches in the x_1 - x_2 plane will deviate from the diagonal $x_1 = -x_2$ by a distance of at most of order $\epsilon\sqrt{1/\epsilon} = \sqrt{\epsilon}$.

5.5. 'Triggering cross' for the second branching

We now discuss the location of the second evolutionary branching. As in the previous section, this location is determined by mutual invasibility analysis, but now for the dimorphic case $D = \{x_1, x_2\}$. The second evolutionary branching may occur on either of the two branches stemming from the first evolutionary branching. We first discuss the condition for the case where the second evolutionary branching occurs on the upper branch of the dimorphic population (an example is shown in Fig. 3a). With $z = x_2 \pm \epsilon$ the corresponding conditions for mutual invasibility are $M_z^{\{x_1, x_2\}} > 1$ and $M_{x_2}^{\{x_1, z\}} > 1$. Analysis of the signs of $M_z^{\{x_1, x_2\}} - 1$ and $M_{x_2}^{\{x_1, z\}} - 1$, reported in Appendix B, shows that in view of Eq. (26) the second evolutionary branching of the upper branch becomes possible in the region

$$x_1 = -x_\alpha + O(\sqrt{\epsilon}), \quad x_2 = x_\alpha + O(\sqrt{\epsilon}), \quad (27)$$

where x_α is given by Eq. (11). More precisely, we show (see Appendix C) that the second evolutionary branching of the upper branch with high probability occurs in the region

$$(x_1 + x_\alpha)c_\alpha - x_2 + x_\alpha = O(\epsilon) \quad (28)$$

which is a neighbourhood of the straight line

$$(x_1 + x_\alpha)c_\alpha = x_2 - x_\alpha. \quad (29)$$

The coefficient c_α is given by

$$c_\alpha = \frac{(1 + 2\alpha)^2 \log(1 + 2\alpha) - 2\alpha(1 + \alpha)}{(1 + 2\alpha)(2\alpha - 1) \log(1 + 2\alpha) + 2\alpha(1 + \alpha)}. \quad (30)$$

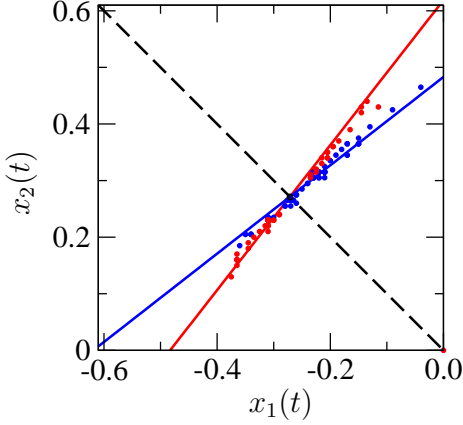


Figure 8: Locations of the second evolutionary branching in the x_1 - x_2 -plane. Shown are the lines (29), solid blue, and (32), solid red, as well as the line corresponding to symmetric evolution of (x_1, x_2) (black dashed). Results of simulations determining the locations of second branchings are shown as points. Blue points correspond to cases where the second branching occurred on the upper branch, red points to branchings of the lower branch. Parameters: $\alpha = 9$, $\epsilon = 10^{-2}$, $N = 4 \times 10^6$, $\mu = 2.5 \times 10^{-10}$, $\theta t \leq 4 \times 10^3$. Six sets of initial conditions were used: $(j_1, j_2) = (-20, 30), (-30, 20), (-10, 35), (-25, 25), (-35, 10)$, and $(-1, 45)$.

It is shown in Appendix A that c_α satisfies $c_\alpha < 1$. The minimum value of c_α equals $0.7732\dots$ and is achieved at $\alpha = 4.0533\dots$

Turning to the possibility that the second branching occurs on the lower branch we find a similar condition

$$(x_2 - x_\alpha)c_\alpha - x_1 - x_\alpha = O(\epsilon) \quad (31)$$

corresponding to an ϵ -neighbourhood of another straight line

$$x_1 + x_\alpha = (x_2 - x_\alpha)c_\alpha. \quad (32)$$

Taking these results together, we have shown that the second evolutionary branching is expected to occur in the vicinity of the lines Eqs. (29) and (32). If the pair (x_1, x_2) of diverging branches comes close to the line Eq. (29), then a second evolutionary branching of the upper branch becomes possible. Conversely, when the pair (x_1, x_2) comes close to the line Eq. (32), a second evolutionary branching may occur on the lower branch. Note that since μ is assumed to be independent of the trait value, our answers Eqs. (29) and (32) depend only on α . The cross formed by the pair of lines (29), and (32) is centered at the point $(-x_\alpha, x_\alpha)$ in the x_1 - x_2 -plane.

Eqs. (29), (32) are consistent with results of direct numerical simulations of the model. Fig. 8 shows the lines (29) and (32) for $\alpha = 9$ and $\epsilon = 10^{-2}$, as well as the locations of second evolutionary branchings for an ensemble of simulations of the stochastic, individual-based model. Six sets of dimorphic initial conditions were used, $(x_1, x_2) = (-0.2, 0.3), (-0.3, 0.2), (-0.25, 0.25), (-0.35, 0.1), (-0.1, 0.35)$ and $(0.01, 0.45)$. We remark that the last three initial conditions deviate substantially from the diagonal (dashed line in Fig. 4). As shown above, paths in the x_1 - x_2 -plane typically exhibit deviations of order $\sqrt{\epsilon}$ from the diagonal. We have nevertheless included dimorphic initial conditions far from the diagonal in

order to clearly exhibit the locations of the second evolutionary branching.

In Fig. 8, blue dots correspond to cases where the second evolutionary branching occurred on the upper branch of the dimorphic state, while red dots correspond to cases where the second evolutionary branching occurred on the lower branch. The coordinates of the dots show the location of the second branching event in the x_1 - x_2 -plane (we determined this location by taking the average of the trait values of the critical state of coexistence prior to the second branching, as well as the coordinate of the second branch when coexistence first occurred). Fig. 8 is consistent with the theoretical expectation that second branchings of the upper branch are triggered by the line Eq. (29), while the second branchings of the lower branch are triggered by the line Eq. (32).

6. Conclusions

We have analysed evolutionary branching in a stochastic, individual-based model for the evolution of a trait value subject to small but non-negligible mutational step sizes ϵ . We found that the branching patterns are in general asymmetric at distinct, non-null mutational steps, complementing the standard adaptive dynamics predictions Geritz et al. (1998) valid in the limit $\epsilon \rightarrow 0$. In particular we found conditions describing the locations where the first branching may occur, Eq. (23), and where the second branching may take place, Eqs. (27), (28), and (31). Results of simulations of a stochastic, individual-based model at small mutational step sizes ϵ were seen to be consistent with these conditions. Our results are derived in the limit of large population sizes N and small mutation rates μ , but allow for fixed mutational step sizes. In this respect our results Eqs. (23), (27), (28), and (31) complement the established approach.

We have also demonstrated that population-size fluctuations in the critical state of coexistence prior to a branching may erase this state. Indeed, its sojourn time depends sensitively on mutational step size. Evolutionary branching occurs typically only when this time is much longer than the time between successful mutations.

Acknowledgments

Financial support from the Swedish Research Council (SS, BM), the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine (BM), by the Bank of Sweden Tercentenary Foundation (SS) is gratefully acknowledged. VV was partially supported by the grant RFBR 11-01-00139 and the program "Dynamical systems and control theory" of the Russian Academy of Sciences.

References

Boettiger C. Stochastic limits to evolutionary branching, 2010. Lab notebook <http://www.carlboettiger.info/archives/715>

Champagnat, N., Ferri Ore, R., and Ben Arous, G. The canonical equation of adaptive dynamics: A mathematical view, 2001. *Selection* 2 71–81

Champagnat, N. A microscopic interpretation for adaptive dynamics trait substitution sequence models, 2006. *Stoch. Proc. Appl.* 116 1127–1160

Champagnat, N. and Méléard, S. Polymorphic evolution sequence and evolutionary branching, 2011. *Probab. Theory Rel. Fields* 151 45–94

Christiansen, F.B. and Loeschke, V. Evolution and intraspecific exploitative competition I. One-locus theory for small additive gene effects, 1980. *Theor. Pop. Biol.* 18 297–313

Collet P., Méléard S., and Metz J. A. J. A rigorous model study of the adaptive dynamics of Mendelian diploids, 2011. <http://arxiv.org/abs/1111.6234>

Dieckmann U. and Law R. The dynamical theory of coevolution: A derivation from stochastic ecological processes, 1996. *J. Math. Biol.* 34 579–612

Diekmann O. A beginner's guide to adaptive dynamics, 2004. *Banach Center Publications* 63, Polish Academy of Sciences

Durinx M., Metz J.A.J., and Meszéna G., Adaptive dynamics for physiologically structured population models, 2008. *J. Math. Biol.* 56 673–742

Geritz, S. A. H., Kisdi É., Meszéna G., and Metz J. A. J. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree, 1998. *Evol. Ecol.* 12 35–57

Haccou, P., Jagers, P., and Vatutin, V. *Branching Processes: Variation, Growth, and Extinction of Populations*, 2005. Cambridge U. Press, Cambridge

Iwasa I., Pomianowski A., and Nee S. The evolution of costly mate preferences. II. The “handicap” principle, 1991. *Evolution* 45 1431–1442

Jagers P. and Klebaner F. Population-size-dependent, age-structured branching processes linger around their carrying capacity, 2011. *J. Appl. Prob.* 48A, 249–262

Klebaner F., Sagitov S., Vatutin V., Haccou P., and Jagers P. Stochasticity in the adaptive dynamics of evolution: the bare bones, 2011. *J. Biol. Dynamics* 5 147–162

Lande R. Natural selection and random genetic drift in phenotypic evolution, 1976. *Evolution* 30 314–334

Méléard, S. and Tran V.C. Trait substitution sequence process and canonical equation for age-structured populations, 2009. *J. Math. Biol.* 58 881–921

Metz J.A.J. Lecture at Mathematical Biosciences Institute (MBI), 2011. Workshop Stochastic Processes in Cell and Population Biology

Metz J.A.J. Adaptive Dynamics, 2012. In: Hastings A. and Gross L. J. (eds.) *Sourcebook in Theoretical Ecology*, Cal. U. Press, 7–17

Metz, J.A.J., S.A.H. Geritz, G. Meszéna, F.J.A. Jacobs, and J.S. van Heerwaarden. Adaptive dynamics: A geometrical study of the consequences of nearly faithful reproduction. In: S.J. van Strien and S.M. Verduyn-Lunel (eds.) *Stochastic and spatial structures of dynamical systems*, 1996. North Holland, Elsevier, 183–231

Su-Chan Park, Simon D., and Krug, J. The speed of evolution in large asexual populations, 2010. *J. Stat. Phys.* 138 381–410

Wakeley, J. *Coalescent Theory: An Introduction*, 2008. Roberts & Company Publishers, Greenwood Village, Colorado.

Waxman D. and Gavrillets S. Issues of terminology, gradient dynamics and the ease of sympatric speciation in Adaptive Dynamics, 2005. *J. Evol. Biol.* 18 1214–1219

Waxman D. and Gavrillets S. 20 Questions on Adaptive Dynamics, 2005. *J. Evol. Biol.* 18 1139–1154

Appendix A

This appendix is divided into three sections summarising material needed in appendices B and C.

1. We discuss the properties of the three functions:

$$C(\alpha) = \frac{\alpha}{2} - ((1 + \alpha)^2 - \alpha^2)x_\alpha^2, \quad (33)$$

$$c_1(\alpha) = \frac{(1 + 2\alpha)^2}{2\alpha(1 + \alpha)} \log(1 + 2\alpha) - 1, \quad (34)$$

$$c_2(\alpha) = \frac{(1 + 2\alpha)(2\alpha - 1)}{2\alpha(1 + \alpha)} \log(1 + 2\alpha) + 1. \quad (35)$$

Here x_α is defined by Eq. (11). We show that $C(\alpha) > 0$, and $0 < c_1(\alpha) < c_2(\alpha)$ for all positive values of α . The latter inequalities

in view of $c_\alpha = c_1(\alpha)/c_2(\alpha)$ imply that the constant c_α , defined by Eq. (30) and determining the slopes of the triggering lines Eqs. (29), (32), satisfies $0 < c_\alpha < 1$.

To see that $C(\alpha) > 0$ we observe that

$$C(\alpha) = \frac{\alpha}{2} - \frac{1 + 2\alpha}{4(1 + \alpha)} \log(1 + 2\alpha) = \frac{h(1 + 2\alpha)}{8(1 + \alpha)}, \quad (36)$$

where $h(x) = x^2 - 2x \log x - 1$. The function $h(x)$ takes positive values for all $x > 1$, since $h(1) = 0$ and $h'(x) > 0$ for $x > 1$. Thus $C(\alpha) > 0$ for all positive α .

Similarly, the fact that $0 < c_1(\alpha) < c_2(\alpha)$ follows from

$$c_1(\alpha) = \frac{h_1(1 + 2\alpha)}{2\alpha(1 + \alpha)}, \quad c_2(\alpha) = c_1(\alpha) + \frac{4}{\alpha}C(\alpha), \quad (37)$$

where $h_1(x) = x^2 \log x - \frac{x^2-1}{2}$ is positive for $x > 1$ (since $h_1(1) = 0$ and $h'_1(x) > 0$).

2. We discuss properties of the equilibrium densities in the dimorphic case $D = \{x_1, x_2\}$. In this case, the equilibrium densities $f_{x_2}^{\{x_1, x_2\}}$ and $f_{x_1}^{\{x_1, x_2\}}$ satisfying the system Eq. (6):

$$\begin{aligned} R_{x_1 x_2} f_{x_2}^{\{x_1, x_2\}} + f_{x_1}^{\{x_1, x_2\}} &= 1, \\ f_{x_2}^{\{x_1, x_2\}} + R_{x_2 x_1} f_{x_1}^{\{x_1, x_2\}} &= 1, \end{aligned}$$

can be represented as $f_{x_2}^{\{x_1, x_2\}} = \phi(x_1, x_2)$ and $f_{x_1}^{\{x_1, x_2\}} = \phi(x_2, x_1)$ with

$$\phi(x_1, x_2) = \frac{1 - R_{x_2 x_1}}{1 - R_{x_1 x_2} R_{x_2 x_1}} = \frac{1 - e^{-(x_2 - x_1)(\alpha x_2 - (2 + \alpha)x_1)}}{1 - e^{-2(1 + \alpha)(x_1 - x_2)^2}}. \quad (38)$$

In particular, in the symmetric case $x_1 = -x$ and $x_2 = x$, we obtain

$$\phi(-x, x) = \phi(x, -x) = \frac{1}{1 + e^{-4(1 + \alpha)x^2}}. \quad (39)$$

3. We quote the Taylor expansion for the function defined by Eq. (5) (recall that $M_x^D = 1$ for $x \in D$)

$$\begin{aligned} M_z^D - 1 &\approx (\alpha x - (1 + \alpha)A_x^D) \Delta \\ &+ \left(\frac{\alpha}{2} - \alpha^2 x^2 + 2\alpha(1 + \alpha)x A_x^D - (1 + \alpha)^2 B_x^D \right) \Delta^2 \end{aligned} \quad (40)$$

valid for small values of $\Delta = z - x$. Here the terms

$$A_x^D = \sum_{y \in D} y R_{xy} f_y^D, \quad B_x^D = \sum_{y \in D} y^2 R_{xy} f_y^D$$

assume the form of first and second moments of the coexisting trait values using the weights $R_{xy} f_y^D$ satisfying condition Eq. (6).

Appendix B

In this appendix we discuss in which region of the x_1 - x_2 -plane the second evolutionary branching becomes possible. It may occur on either of the two branches stemming from the first evolutionary branching. We start by discussing the evolution of the upper branch of the dimorphic population on its way

towards the second branching (an example is shown in Fig. 3a). With $z = x_2 \pm \epsilon$ the corresponding conditions for mutual invasibility are $M_z^{[x_1, x_2]} > 1$ and $M_{x_2}^{[x_1, z]} > 1$. The signs of $M_z^{[x_1, x_2]} - 1$ and $M_{x_2}^{[x_1, z]} - 1$ are determined by Eq. (40) with

$$\begin{aligned} A_{x_2}^{[x_1, x_2]} &= x_2 \phi(x_1, x_2) + x_1 R_{x_2 x_1} \phi(x_2, x_1) \\ &= x_1 + (x_2 - x_1) \phi(x_1, x_2), \end{aligned} \quad (41)$$

$$\begin{aligned} B_{x_2}^{[x_1, x_2]} &= x_2^2 \phi(x_1, x_2) + x_1^2 R_{x_2 x_1} \phi(x_2, x_1) \\ &= x_1^2 + (x_2^2 - x_1^2) \phi(x_1, x_2), \end{aligned} \quad (42)$$

where $\phi(x_1, x_2)$ is given by Eq. (38). Using Eqs. (40), (41), and (42) we find

$$M_z^{[x_1, x_2]} - 1 \approx C_1(x_1, x_2)(z - x_2) + C_2(x_1, x_2)\epsilon^2, \quad (43)$$

where

$$\begin{aligned} C_1(x_1, x_2) &= \alpha x_2 - (1 + \alpha)x_1 \\ &\quad - (1 + \alpha)(x_2 - x_1)\phi(x_1, x_2), \end{aligned} \quad (44)$$

$$\begin{aligned} C_2(x_1, x_2) &= \frac{\alpha}{2} - \{\alpha x_2 - (1 + \alpha)x_1\}^2 \\ &\quad + (1 + \alpha)(x_2 - x_1)\{(\alpha - 1)x_2 \\ &\quad - (1 + \alpha)x_1\}\phi(x_1, x_2). \end{aligned} \quad (45)$$

Using a counterpart of Eq. (16) for the survival probability $P_{x_2+\epsilon}(x_1, x_2)$ of a single $(x_2 + \epsilon)$ -mutant in a stable (x_1, x_2) -dimorphic population we find that

$$P_{x_2+\epsilon}(x_1, x_2) \approx 2\epsilon(x_2 - x_1)(\alpha - (x_2 - x_1)^{-1}x_1 - (1 + \alpha)\phi(x_1, x_2)). \quad (46)$$

Similarly, consider the possibility that the second branching occurs on the lower branch. In this case we have for $z = x_1 \pm \epsilon$

$$M_z^{[x_1, x_2]} - 1 \approx C_1(x_2, x_1)(z - x_1) + C_2(x_2, x_1)\epsilon^2 \quad (47)$$

which yields

$$\begin{aligned} P_{x_1-\epsilon}(x_1, x_2) &\approx 2\epsilon(x_2 - x_1)(\alpha + (x_2 - x_1)^{-1}x_2 \\ &\quad - (1 + \alpha)\phi(x_2, x_1)). \end{aligned} \quad (48)$$

Close inspection of Eqs. (46) and (48) for $x_1 < 0 < x_2$ using Eq. (38) reveals that

$$P_{x_1-\epsilon}(x_1, x_2) > P_{x_2+\epsilon}(x_1, x_2) \text{ given } |x_1| < x_2. \quad (49)$$

As explained in Section 5.4, this is one of the factors ensuring the negative dependence among the steps of the random walk in the x_1 - x_2 -plane leading to the condition Eq. (26).

Applying Eq. (26) we can further refine our previous analysis and deduce from Eq. (43)

$$M_z^{[x_1, x_2]} - 1 \approx C_1(x_1, x_2)(z - x_2) + C_2(-x_2, x_2)\epsilon^2. \quad (50)$$

Replacing x_1 by $-x_2$ in this expression incurs error of order $\sqrt{\epsilon}$. This implies:

$$M_z^{[x_1, x_2]} - 1 = C_1(-x_2, x_2)(z - x_2) + O(\epsilon^{3/2}). \quad (51)$$

It follows from Eq. (39) that

$$C_1(-x, x) = (1 + 2\alpha)x - \frac{2(1 + \alpha)x}{1 + e^{-4(1+\alpha)x^2}}, \quad (52)$$

$$C_2(-x, x) = \frac{\alpha}{2} - (1 + 2\alpha)^2 x^2 + \frac{4\alpha(1 + \alpha)x^2}{1 + e^{-4(1+\alpha)x^2}}. \quad (53)$$

Now let x_α be given by Eq. (11). Since $C_1(-x, x)$ is positive for $x \in (0, x_\alpha)$, Eq. (51) confirms that the replacement process on the upper branch goes upward until $x_2 = x_\alpha + O(\sqrt{\epsilon})$.

Similarly, consider the possibility that the second branching occurs on the lower branch. In this case using Eqs. (26) and (47) we get a lower branch counterpart of Eq. (51)

$$M_z^{[x_1, x_2]} - 1 = C_1(x_1, -x_1)(x_1 - z) + O(\epsilon^{3/2}) \quad (54)$$

implying that the replacement on the lower branch goes downward until $x_1 = -x_\alpha + O(\sqrt{\epsilon})$. We conclude that the second evolutionary branching on one of the two branches is only possible in the region Eq. (27).

Appendix C

In this appendix we establish Eq. (28) in Section 5.5 which in turn implies Eq. (29) determining the triggering line for the upper branch. Given Eq. (27), the deviations $\delta_1 = x_1 + x_\alpha$ and $\delta_2 = x_2 - x_\alpha$ satisfy $\delta_i = O(\sqrt{\epsilon})$. We demonstrate first that

$$\begin{aligned} C_1(x_1, x_2) &= \frac{1}{2}c_1(\alpha)\delta_1 - \frac{1}{2}c_2(\alpha)\delta_2 + c_{11}(\alpha)\delta_1^2 \\ &\quad + c_{12}(\alpha)\delta_1\delta_2 \\ &\quad + c_{22}(\alpha)\delta_2^2 + o(\epsilon), \end{aligned} \quad (55)$$

where $c_i(\alpha)$ are given by (34) and (35), while the exact form of the constants $c_{ij}(\alpha)$ does not matter. Using $\phi(-x_\alpha, x_\alpha) = \frac{1+2\alpha}{2(1+\alpha)}$ and the Taylor expansion of Eq. (38) we obtain

$$\begin{aligned} \phi(x_1, x_2) &= \frac{1 + 2\alpha}{2(1 + \alpha)} + \delta_1\phi_1 + \delta_2\phi_2 + \frac{\delta_1^2}{2}\phi_{11} \\ &\quad + \frac{\delta_1\delta_2}{2}\phi_{12} + \frac{\delta_2^2}{2}\phi_{22} + o(\epsilon), \end{aligned} \quad (56)$$

where $\phi_i = \frac{\partial}{\partial x_i}\phi(x_1, x_2)|_{x_1=-x_\alpha, x_2=x_\alpha}$ and ϕ_{ij} are the corresponding second order derivatives. It follows in view of Eq. (45), that Eq. (56) holds with

$$\begin{aligned} c_1(\alpha) &= -1 - 4(1 + \alpha)x_\alpha\phi_1, \\ c_2(\alpha) &= 1 + 4(1 + \alpha)x_\alpha\phi_2, \\ c_{11}(\alpha) &= (1 + \alpha)(\phi_1 - x_\alpha\phi_{11}), \\ c_{12}(\alpha) &= (1 + \alpha)(\phi_2 - \phi_1 - x_\alpha\phi_{12}), \\ c_{22}(\alpha) &= -(1 + \alpha)(\phi_2 + x_\alpha\phi_{22}), \end{aligned} \quad (57)$$

so that it remains only to verify that the new relations for $c_i(\alpha)$ agree with Eqs. (34) and (35). For this it is sufficient to note that $\phi_1 = -\frac{(1+2\alpha)^2 x_\alpha}{2\alpha(1+\alpha)}$ and $\phi_2 = \frac{(1+2\alpha)(2\alpha-1)x_\alpha}{2\alpha(1+\alpha)}$. This follows from Eq. (38). We conclude that, given Eq. (27), we have

$$\begin{aligned} 2C_1(x_1, x_2) &= c_1(\alpha)(x_1 + x_\alpha) \\ &\quad - c_2(\alpha)(x_2 - x_\alpha) + O(\epsilon). \end{aligned} \quad (58)$$

To conclude our derivation of Eq. (28) we notice that in accordance with Eqs. (50) and (27) we have

$$M_z^{\{x_1, x_2\}} - 1 \approx C_1(x_1, x_2)(z - x_2) + C(\alpha)\epsilon^2, \quad (59)$$

where $C(\alpha) = C_2(-x_\alpha, x_\alpha)$ is a positive constant given by Eq. (33). Eq. (51) implies that the inequality $M_{x_2+\epsilon}^{\{x_1, x_2\}} > 1$ is equivalent to $C_1(x_1, x_2) > -\epsilon C(\alpha)$. Similarly, due to Eq. (51), the inequality $M_{x_2}^{\{x_1, x_2+\epsilon\}} > 1$ translates into $C_1(x_1, x_2) < \epsilon C(\alpha)$. Combining the last two relations with Eq. (58) we derive Eq. (28). In other words, second evolutionary branching on the upper branch of the dimorphic population occurs in an ϵ -neighbourhood of the line given by Eq. (29).

Using Eq. (47) and repeating the arguments summarised above, we find the condition (31) for the second evolutionary branching to occur on the lower branch of the dimorphic population. This branching occurs in an ϵ -neighbourhood of the straight line Eq. (32).